

# Analysis of feedback systems with structured uncertainties

John Doyle

*Indexing terms:* Control theory, Feedback, Robustness, Sensitivity

**Abstract:** The paper introduces a general approach for analysing linear systems with structured uncertainty based on a new generalised spectral theory for matrices. The results of the paper naturally extend techniques based on singular values and eliminate their most serious difficulties.

## 1 Introduction

The last several years has seen something of a revolution in multivariable control theory. Central to this revolution is a renewed focus on the fundamental issue in feedback design: providing performance in the face of uncertainty. This development has been supported by a renewed interest in the frequency domain. An indication of the growing acceptance of these points of view can be found in the recent Special Issue on 'Linear multivariable control systems' in the February 1981 *IEEE Transactions on Automatic Control*, where five of the first six articles [1–5] deal with uncertainty, robustness and sensitivity using frequency-domain tools.

There are many results of both practical and theoretical importance which have come out of this research. As a consequence, it is now well known that multivariable sensitivity and robustness cannot be reliably evaluated one loop at a time [1, 6]. Singular-value methods have proved useful in providing both a reliable multivariable measure of sensitivity and robustness [1, 3, 7] and, more generally, a framework within which to develop reliable multiloop versions of the main tools of classical control theory [1–9].

While these results are reasonably well understood within the control theory community, many of the more subtle aspects of multivariable feedback systems are less well understood. Multiloop systems have important properties that have no analogues in the single-loop case. For example, there are problems caused by the fact that the signals and responses in multiloop systems vary not only with frequency, but also with direction. While some published results [8–10] have given preliminary indications of this, the issue of directionality in multiloop systems is largely unexplored.

The most serious limitation associated with the methods based on singular values is that they treat a limited, albeit important, class of uncertainties. Roughly speaking, most of these results give, at best, exact characterisation (i.e. involving necessary and sufficient conditions) of the feedback properties of a system model relative to perturbations which are norm bounded but otherwise unconstrained. There is no systematic mechanism for exploiting information about the structure of a perturbation.

The aim of this paper is to develop techniques for analysing systems with arbitrary constraints on their structure. Section 2 discusses some particularly important examples of structured uncertainties in an attempt to motivate the general results in later Sections.

Sections 3–5 present the main theoretical results. In Section 3, the block-diagonal perturbation problem is formulated. This problem is quite general since any norm-bounded perturbation problem, regardless of structure, can be trivially rewritten as a block-diagonal perturbation problem. This pro-

blem leads to the definition of a function  $\mu$  that provides necessary and sufficient conditions for structured matrix perturbation problems. Section 4 examines the properties of the function  $\mu$  and expresses it in familiar matrix algebraic terms. Section 5 develops the tools necessary to find gradients for singular values, which are used in Section 6 to provide techniques to compute  $\mu$  in important special cases. Section 7 has some examples and a discussion of recent experience in computing  $\mu$ , and Section 8 gives a summary and conclusions.

## 2 Importance of the block-diagonal perturbation problem

Analysis methods based on singular values have been successful in providing a framework within which to develop multiloop generalisations of classical single-loop techniques. While this has been an important development, singular-value methods have serious limitations. Consider, for example, the problem of analysing a linear multivariable feedback system with two multiplicative perturbations appearing simultaneously at the inputs and outputs [8–10].

There are essentially two direct applications of singular-value techniques to this problem. First, the system may be rewritten to isolate the two perturbations as a single perturbation with a two-block-diagonal structure. This structure can then be ignored by replacing the block-diagonal perturbation with one full matrix perturbation. In this configuration, standard singular-value analysis can be applied. Of course, the results may be, in general, arbitrarily conservative. A second approach is simply to treat the two perturbations one at a time, leading possibly to arbitrarily optimistic answers.

If these first two approaches to the simultaneous perturbation problem happen to give the same results, then the 'true' answer is obviously determined. In general, this will not be the case. It should be emphasised again that these two approaches can yield arbitrarily bad estimates; i.e., depending on the problem, the gap between the upper and lower bounds may be arbitrarily large. No other information can be obtained from these bounds other than that the 'true' answer lies between them. This will become evident in the course of this paper.

An extension of the direct application of singular values is to analyse the differential sensitivity of the singular values evaluated at one point relative to perturbations at the other [10]. Reference 10 is extremely well written and would provide excellent background for the present paper. However, this approach suffers from the fundamental difficulty associated with differential sensitivity techniques in not applying to large perturbations. In addition, the results in Reference 10 provided only directional sensitivity information (i.e. directional derivatives) when singular values are not clustered. If differential sensitivity methods are to be effective, it is often important that worst-case directions (as with gradients) be obtained. The results of this paper can be modified for computing worst-case differential directions for clustered singular values. Although not developed further here, this may make it

possible to complete the results in Reference 10 for clustered singular values.

What is really desired, of course, is an exact analysis of the original simultaneous independent perturbation problem, which applies to large perturbations. One application of the results in this paper is just such an analysis method. This will be illustrated in Section 7 by a simple 'textbook' example.

The problem of evaluating robustness with respect to simultaneous input and output perturbations is a special case of the more general problem of evaluating robustness with respect to perturbations with arbitrary constraints on their structure. Consider the general problem of an interconnected linear system with multiple independent norm-bounded perturbations occurring throughout. By rearranging the system, it is always possible to isolate the perturbations as a single large block-diagonal perturbation.

Stability analysis, for example, then essentially boils down to ensuring that  $I + M\Delta$  remains nonsingular at all frequencies and for all  $\Delta$  under consideration. Here,  $\Delta$  is a block-diagonal perturbation and  $M$  is the transfer function from the collective outputs of the perturbations to their inputs. Readers unfamiliar with manipulating systems to isolate specific components as a diagonal matrix are encouraged to try some examples and convince themselves that it can be done. This notion is essential to understanding how the results in this paper may be applied to analyse control systems.

The key problem then involves the matrix problem of determining necessary and sufficient conditions such that  $\det(I + M\Delta) \neq 0$ . All norm-bounded linear perturbation problems reduce almost trivially to this. The main results in this paper are a partial solution to the general block-diagonal perturbation problem and a complete solution for the case of three or fewer blocks.

### 3 Definitions and problem formulation

The following notation will be used throughout:

$\mathcal{M}(k)$  = algebra of complex  $k \times k$  matrices  
 $\mathcal{U}(k)$  = unitary matrices in  $\mathcal{M}(k)$   
 $\sigma_{\max}(M)$  = maximum singular value of  $M$   
 $\rho(M)$  = spectral radius = magnitude of largest eigenvalue  
 $\lambda_{\min}(H)$  = least (most negative) eigenvalue for Hermitian  $H$   
 $M^*$  = conjugate transpose of  $M$   
 $\text{diag}(M_1, M_2, M_3, \dots, M_n)$  = block-diagonal matrix with  $M_j$  (not necessarily square) on the diagonal

To provide a description of block-diagonal perturbations, let  $\mathcal{K} = (m_1, m_2, \dots, m_n, k_1, k_2, \dots, k_n)$  be a  $2n$ -tuple of positive integers. All the definitions that follow depend on

$$\mathcal{D} = \{\text{diag}(d_1 I_{k_1}, d_2 I_{k_2}, \dots, d_m I_{k_m}) \mid d_i \in \mathbb{R}^+ = (0, \infty)\}$$

$$\mathcal{S} = \{\text{diag}(d_1, d_2, \dots, d_K) \mid d_i \in [0, \infty)\}$$

$$\mathcal{V} = \{\text{diag}(\overbrace{v_1, \dots, v_1}^{m_1}, \overbrace{v_2, \dots, v_2}^{m_2}, \dots, \overbrace{v_n, \dots, v_n}^{m_n}) \mid v_j \in \mathbb{C}^{k_j \times 1}, v_j^* v_j = 1, j = 1, 2, \dots, n\}$$

$$\mathcal{V} \subset \mathbb{C}^{K \times m}$$

$\mathcal{K}$ , but, to simplify notation, this dependency will not be explicitly represented. Unless specifically noted otherwise,  $\mathcal{K}$  will be assumed to be an arbitrary but fixed  $2n$ -tuple of positive integers. Let

$$K = \sum_{j=1}^n m_j k_j \quad \text{and} \quad m = \sum_{j=1}^n m_j$$

For each  $\delta \in \mathbb{R}, \delta \geq 0$ , let  $X_\delta \subset \mathcal{M}(K)$  be

$$X_\delta = \{\text{diag}(\overbrace{\Delta_1, \dots, \Delta_1}^{m_1}, \overbrace{\Delta_2, \dots, \Delta_2}^{m_2}, \Delta_3, \dots, \Delta_{n-1}, \overbrace{\Delta_n, \dots, \Delta_n}^{m_n}) \mid \Delta_j \in \mathcal{M}(k_j) \text{ and } \sigma_{\max}(\Delta_j) \leq \delta \text{ for each } j = 1, 2, \dots, n\}$$

Let

$$X_\infty = \bigcup_{j=1}^{\infty} X_j$$

Thus  $X_\delta$  is the set of block-diagonal matrices with structure determined by  $\mathcal{K}$  whose norm is not greater than  $\delta$ , and  $X_\infty$  is all such matrices with no restriction on the norm.

What is desired is a function (depending, of course, on  $\mathcal{K}$ )

$$\mu: \mathcal{M}(K) \rightarrow [0, \infty)$$

with the property that for all  $M \in \mathcal{M}(K)$

$$\det(I + M\Delta) \neq 0 \quad \text{for all } \Delta \in X_\delta \quad (1)$$

if, and only if,

$$\delta \mu(M) < 1$$

This could be taken as a definition of  $\mu$ . Alternatively,  $\mu$  could be defined as

$$\mu(M) = \begin{cases} 0 & \text{if no } \Delta \in X_\infty \text{ solves } \det(I + M\Delta) = 0 \\ \left( \min_{\Delta \in X_\infty} \left\{ \sigma_{\max}(\Delta) \mid \det(I + M\Delta) = 0 \right\} \right)^{-1} & \text{otherwise} \end{cases}$$

This definition of  $\mu$  shows that a well defined function satisfies eqn. 1. It probably has little additional value since the optimisation problem involved does not appear to have any useful properties.

In order to proceed further, some additional definitions are needed. Again, assume the  $2n$ -tuple  $\mathcal{K}$  is given.

Then let

$$\mathcal{U} = \mathcal{U}(K) \cap X_1 \quad (\text{block-diagonal unitary matrices})$$

Note that these definitions all depend on  $\mathcal{K}$ , but to explicitly represent the dependency (for example,  $\mu_{\mathcal{K}}, X_{\mathcal{K}, \delta}, \mathcal{U}_{\mathcal{K}}$  etc.) would be unnecessarily cumbersome. Using these definitions, several useful properties of  $\mu$  are stated in the following Section, leading up to the first main result of the paper.

The problem formulation taken here is certainly not the most general possible. For example, the entire development in

this paper could be done with nonsquare perturbations with no conceptual change in the results or proofs. The notation, however, would become even more cumbersome than it already is. The particular approach taken is the simplest that allows both  $\rho$  and  $\sigma_{\max}$  to emerge as special cases of  $\mu$ . Hopefully, this will aid the reader by maintaining continuity with standard linear algebra as well as with some of the more popular existing methods in multivariable control.

#### 4 Properties of $\mu$

The following properties of  $\mu$  are easily proven from eqn. 1. As always,  $\mathcal{K} = (m_1, \dots, m_n, k_1, \dots, k_n)$  is an arbitrary  $2n$ -tuple for some  $n \in \mathbb{Z}^+$ , unless specified otherwise, and

$$K = \sum_{j=1}^n m_j k_j$$

- (a)  $\mu(\alpha M) = |\alpha| \mu(M)$  for all  $M \in \mathcal{M}(K)$
- (b)  $\mu(I) = 1$
- (c)  $\mu(AB) \leq \sigma_{\max}(A) \mu(B)$  for all  $A, B \in \mathcal{M}(K)$
- (d)  $\mu(\Delta) = \sigma_{\max}(\Delta)$  for all  $\Delta \in X_\delta$
- (e) If  $n = 1$  and  $m_1 = 1$ , then

$$\mu(M) = \sigma_{\max}(M) \quad \text{for all } M \in \mathcal{M}(K)$$

- (f) If  $n = 1, k_1 = 1$ , then

$$K = m_1$$

$$X_\delta = \{\lambda I \mid \lambda \in \mathbb{C}, |\lambda| \leq \delta\}$$

$$\mu(M) = \rho(M) \quad \text{for all } M \in \mathcal{M}(K)$$

- (g) If  $\Delta \in X_\delta, U \in \mathcal{U}$ , then  $U\Delta \in X_\delta$  and  $\Delta U \in X_\delta$

- (h) For all  $\Delta \in X_\infty$  and for all  $D \in \mathcal{D}$ ,

$$D\Delta D^{-1} = \Delta$$

- (i) For all  $U \in \mathcal{U}$  and  $M \in \mathcal{M}(K)$ ,

$$\mu(MU) = \mu(UM) = \mu(M)$$

- (j) For all  $D \in \mathcal{D}$  and  $M \in \mathcal{M}(K)$ ,

$$\mu(DMD^{-1}) = \mu(M)$$

- (k)  $\max_{U \in \mathcal{U}} \rho(UM) \leq \mu(M) \leq \inf_{D \in \mathcal{D}} \sigma_{\max}(DMD^{-1})$

$$\text{for all } M \in \mathcal{M}(K)$$

- (l) For all  $\Delta \in X_\infty$ , there exist  $U, V \in \mathcal{U}$  and  $\Sigma \in \mathcal{S}$  such that

$$\Delta = U\Sigma V^*$$

- (m)  $\mu$  is continuous in the usual metric topology on  $\mathcal{M}(K)$ .

Property (l) is just a block version of the singular-value decomposition (SVD). Properties (e) and (f) show that  $\mu$  has as special cases both the spectral radius and maximum singular value or spectral norm. Property (i) means that  $\mu$  is  $\mathcal{U}$ -invariant. The main theorem of this Section will show that the left-hand inequality in property (k) is actually an equality and thus that  $\mu$  is, in some sense, the fundamental  $\mathcal{U}$ -invariant. In order to prove this theorem, some simple lemmas involving polynomial equations are needed.

Suppose  $p: \mathbb{C}^k \rightarrow \mathbb{C}$  is a polynomial in  $k$  complex variables of degree no more than  $q \in \mathbb{Z}$  for each variable. For  $z = (z_1, \dots, z_k) \in \mathbb{C}^k$ , let  $\|z\|_\infty = \max |z_j|$  be the usual  $\|\cdot\|_\infty$

norm on  $\mathbb{C}^k$ . Let  $\hat{z} \in \mathbb{C}^k$  be such that  $p(\hat{z}) = 0$  and  $|\hat{z}| = \min_{z \in \mathbb{C}^k} (\|z\|_\infty \mid p(z) = 0)$ ; i.e.  $\hat{z}$  is a solution of  $p(z) = 0$  with minimum  $\|\cdot\|_\infty$  norm.

##### Lemma 1

There exists  $x \in \mathbb{C}^k$  such that  $p(x) = 0$  and  $|x_j| = \|\hat{z}\|_\infty$  for all  $j = 1, \dots, k$ .

##### Proof

Let  $x_j = \hat{z}_j$  for each  $\hat{z}_j$  such that  $|\hat{z}_j| = \|\hat{z}\|_\infty$ . If  $|\hat{z}_j| = \|\hat{z}\|_\infty$  for all  $j$ , then  $x = \hat{z}$  satisfies the lemma. So suppose that one of the components of  $\hat{z}$ , say for convenience  $\hat{z}_k$ , has  $|\hat{z}_k| < \|\hat{z}\|_\infty$ . Then  $p(z) = \sum p_r(z_1, \dots, z_{k-1})z_k^r$ , where each  $p_r$  is a polynomial in  $k-1$  variables. Suppose that for some  $r$ ,  $p_r(\hat{z}_1, \dots, \hat{z}_{k-1}) \neq 0$ . This will lead to a contradiction. By continuity of the roots of a polynomial, for any  $\epsilon > 0$  there exists  $\delta > 0$  such that, for each  $y \in \mathbb{C}^{k-1}$  such that  $\|y - (\hat{z}_1, \dots, \hat{z}_{k-1})\| < \delta$ , there exists a  $w \in \mathbb{C}$  such that  $|w - \hat{z}_k| < \epsilon$  and  $\sum p_r(y)w^r = 0$ . Then, by choosing  $\epsilon$  sufficiently small, there exists a  $\hat{y} \in \mathbb{C}^{k-1}$  and  $\hat{w} \in \mathbb{C}$  such that  $p(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_{k-1}, \hat{w}) = 0$  and  $\|(\hat{y}_1, \dots, \hat{y}_{k-1}, \hat{w})\|_\infty < \|\hat{z}\|_\infty$ . This contradicts the minimising property of  $\hat{z}$ . Thus  $p_r(\hat{z}_1, \dots, \hat{z}_{k-1}) = 0$  for all  $r$ ,  $0 \leq r \leq q$ , and  $p(\hat{z}_1, \dots, \hat{z}_{k-1}, w) \equiv 0$  independent of  $w$ . Let  $x_k = \|\hat{z}\|_\infty$ . This argument may be repeated for each  $j = 1, 2, \dots, k$  so that either  $x_j = \hat{z}_j$  or  $x_j = \|\hat{z}\|_\infty$  and  $p(x) = 0$  for  $x = (x_1, \dots, x_k)$ . This completes the proof.

This lemma may seem a bit obscure at first glance. What it basically says is that if  $\|z\|_\infty$  is to be minimised subject to the constraint that  $p(z) = 0$ , then at least one minimising solution  $x$  lies on the polydisc where each  $|x_j| = \|x\|_\infty$ . This will allow a great simplification in the characterisation of  $\mu$ . First though, note that the proof of lemma 1 immediately implies the following two additional lemmas.

##### Lemma 2

If  $\hat{z}$  is real  $[\text{Im}(\hat{z}_j) = 0 \text{ for all } j]$ , then there exists  $x \in \mathbb{R}^k$  such that  $p(x) = 0$  and  $|x_j| = \|\hat{z}\|_\infty$  for all  $j = 1, 2, \dots, k$ .

##### Lemma 3

If  $\hat{z}$  is real and non-negative  $[\text{Im}(\hat{z}_j) = 0 \text{ and } \text{Re}(\hat{z}_j) \geq 0 \text{ for all } j]$ , then there exists  $x \in \mathbb{R}^k$  non-negative such that  $p(x) = 0$  and  $x_j = \|\hat{z}\|_\infty$ .

##### Theorem 1

For all  $M \in \mathcal{M}(K)$ ,

$$\mu(M) = \max_{U \in \mathcal{U}} \rho(MU)$$

##### Proof

Suppose  $M \in \mathcal{M}(K)$ . If  $\mu(M) = 0$ , then the result follows immediately from property (k), and so assume  $\mu(M) = 1/\delta > 0$ . Then there exists  $\Delta \in X_\delta$  (not necessarily unique) such that  $\sigma_{\max}(\Delta) = \delta$  and  $\det(I + M\Delta) = 0$ . By property (l), there exist  $U, V \in \mathcal{U}$  and  $\Sigma \in \mathcal{S}$  such that

$$\Delta = U\Sigma V^*$$

Then

$$\det(I + M\Delta) = 0$$

if, and only if,

$$\det(I + MU\Sigma V^*) = 0$$

This last equation may be viewed as a polynomial in the diagonal elements of  $\Sigma$ . By assumption,  $\Sigma$  is a minimum norm (now viewed as  $\|\cdot\|_\infty$  on the diagonal elements) solution to this polynomial. This satisfies the conditions of lemma 3 and

therefore  $\Sigma$  may be replaced by a scalar  $\delta I$ . Thus

$$\det(I + MU\Sigma V^*) = 0$$

implies

$$\det(I + \delta MUV^*) = 0$$

which implies

$$\rho(MUV^*) \geq 1/\delta = \mu(M)$$

By property (k), the reverse inequality also holds, and so

$$\max_{U \in \mathcal{U}} \rho(MU) = \mu(M)$$

*Corollary*

$$\mu(M) = \max_{V, W \in \mathcal{V}} \rho(V^* M W) \quad \text{for all } M \in \mathcal{M}(K)$$

*Proof*

$\max_{U \in \mathcal{U}} \rho(MU) = \lambda \Rightarrow$  there exist  $U \in \mathcal{U}$  and  $x \in \mathbb{C}^k$  such that  $MUx = \lambda x$ . This implies there exist

$$V, W \in \mathcal{V} \text{ such that } MWV^*x = \lambda x$$

which implies

$$\rho(MWV^*) = \rho(V^* M W) \geq \lambda$$

which in turn implies

$$\max_{V, W \in \mathcal{V}} \rho(V^* M W) \geq \lambda$$

Similarly,

$$\max_{V, W \in \mathcal{V}} \rho(V^* M W) = \lambda$$

implies

$$\max_{U \in \mathcal{U}} \rho(MU) \geq \lambda$$

Thus

$$\mu(M) = \max_{V, W \in \mathcal{V}} \rho(V^* M W) \quad \text{for all } M \in \mathcal{M}(K)$$

This theorem and corollary express  $\mu$  in terms of familiar linear algebraic quantities. Note that it is now possible to use the corollary to define a general matrix decomposition that would have both the singular-value decomposition and Jordan form as special cases. While such a unification may prove to be of a great theoretical interest, it is not essential to the aims of this paper and will not be pursued further here.

Instead, the remainder of this paper will be devoted to developing methods to compute  $\mu(M)$ . Unfortunately, the optimisation problem expressed in the corollary cannot be solved in general by simple gradient or local techniques. Examples have been generated which have multiple local maxima. While recent computational experience with some alternative algorithms is very promising, no efficient algorithms have been developed which have proven guaranteed convergence to the global maximum.

An alternative approach will now be taken by considering the right-hand side inequality of property (k). In important special cases it will be shown that this inequality is actually an equality, and that  $\inf \sigma_{\max}(DMD^{-1})$  has no local minimum which is not global. In fact, a tedious but straightforward calculation shows that  $\sigma_{\max}(DMD^{-1})$  is convex in  $D$ . The proof of this is omitted since it is not actually needed for any of the main results, but it makes this optimisation problem an attractive

alternative. The first step needed for solving  $\inf \sigma_{\max}(DMD^{-1})$  is a technique for computing descent directions. This problem is taken up in the following Section.

## 5 Differentiability properties of singular values

This Section will develop the necessary tools for computing 'gradients' for singular values. Although singular values are not in general differentiable functions of the matrix elements, it is always possible to compute a generalised gradient which serves the same purpose as would a gradient. This Section will be brief and not at all self-contained and only original results will be proved. An excellent background for this Section is Reference 10, as well as standard texts [11–13].

Suppose  $M: \Omega \rightarrow \mathcal{M}(k)$  is a matrix-valued function whose elements are real analytic in some neighbourhood  $\Omega$  of  $0 \in \mathbb{R}^q$ . Denote by  $\sigma_{\max}(x)$  [for  $\sigma_{\max}(M(x))$ ,  $x \in \Omega$ ] the maximum singular value of  $M$  as a function of  $x \in \Omega \subset \mathbb{R}^q$ .

The directional derivatives of  $\sigma_{\max}(x)$  at  $x = 0$  depend only on  $M(x)$  to first order, and so without loss of generality assume

$$M(x) = M_0 + \sum_{j=1}^q x_j M_j$$

for  $x = (x_1, x_2, \dots, x_q)$ . Using the SVD,  $M(0)$  may be written as

$$M(0) = M_0 = \sigma_{\max}(0) U_1 V_1^* + U_2 \Sigma_2 V_2^*$$

Here  $U_1, V_1 \in \mathbb{C}^{k \times r}$  with  $r$  the multiplicity of  $\sigma_{\max}(0)$  and

$$[U_1 U_2], [V_1 V_2] \in U(k)$$

Suppose  $\sigma_{\max}(0) > 0$  [i.e.  $M(0) \neq 0$ ]. Then, letting  $(\sigma_{\max})_0 = \sigma_{\max}(0)$ ,  $\sigma_{\max}(x)$  may be written as

$$\sigma_{\max}(x) = |(\sigma_{\max})_0 + \lambda_{\max}(\Sigma x_j H_j) + o(x)| \quad (2)$$

where

$$H_j = H_j^* = \text{Re}(U_1^* M_j V_1) = \frac{1}{2}(U_1^* M_j V_1 + V_1^* M_j^* U_1)$$

Define  $\nabla_1 \subset \mathbb{R}^q$  by

$$\nabla_1 = \{x \in \mathbb{R}^q, x_j = v^* H_j v / v^* v = 1 \text{ and } v \text{ is an eigenvector of } \Sigma y_j H_j \text{ for some } 0 \neq y \in \mathbb{R}^q\}$$

and denote by  $\text{co}\nabla_1$  the convex hull of  $\nabla_1$ . Denote by  $\hat{x} = \min(\text{co}\nabla_1)$  the unique point  $\hat{x} \in \text{co}\nabla_1$  that minimises

$$\min_{x \in \text{co}\nabla_1} \|x\|_2$$

A well known property of convex sets that will be used in the following theorem is that  $\langle \hat{x}, x \rangle \geq \|x\|^2$  for all  $x \in \text{co}\nabla_1$  [14]. The significance of  $\nabla_1$  is expressed in the following theorem.

### Theorem 2

If  $\hat{x} \neq 0$ , then there exists  $\epsilon_0 > 0$  such that

$$\sigma_{\max}(-\epsilon \hat{x}) < (\sigma_{\max})_0 \quad \text{for all } 0 < \epsilon < \epsilon_0$$

*Proof*

Let  $v$  be a unit eigenvector associated with the eigenvalue  $\lambda_{\min}(\Sigma \hat{x}_j H_j)$  and let  $y \in \nabla_1$  be defined by  $y_j = v^* H_j v$  for  $j = 1, 2, \dots, q$ . Then

$$\lambda_{\min}(\Sigma \hat{x}_j H_j) = v^* (\Sigma \hat{x}_j H_j) v = \langle \hat{x}, y \rangle > 0$$

Then from eqn. 2 there exists  $\epsilon_0 > 0$  such that  $\sigma_{\max}(-\epsilon \hat{x}) < (\sigma_{\max})_0$  for all  $0 < \epsilon < \epsilon_0$ .

This theorem implies that, although  $\sigma_{\max}(x)$  is not necessarily

differentiable,  $\hat{x} = \min(\text{co}\nabla_1)$  serve one of the purposes of a gradient in providing a descent direction. If  $x \neq 0$ , then a direction (for example,  $-x$ ) can be found which reduces all the singular values in the cluster. In the special case where  $\sigma_{\max}(0)$  has multiplicity one, then  $\nabla_1$  trivially reduces to the ordinary gradient. It may be thought of as the direct limit of the sets of possible gradients in neighbourhoods of 0. Unfortunately,  $\nabla_1$  is a rather awkward set to work with since it depends, in part, on singular vectors of  $M(x)$  for  $x \neq 0$ .

As an alternative, consider the set

$$\nabla_2 = \{x \in \mathbb{R}^q, x_j = v^* H_j v \mid v \in \mathbb{C}^q, v^* v = 1\}$$

Clearly,  $\nabla_1 \subset \nabla_2$ . While  $\nabla_2$  has a much simpler description than  $\nabla_1$ , it serves the same purpose, as is shown by the following lemma and theorem.

**Lemma 4**

$$\min(\text{co}\nabla_1) = \min(\text{co}\nabla_2)$$

**Proof**

Let  $y \in \nabla_2$  and  $\hat{x} = \min(\text{co}\nabla_1)$ . Then  $y_j = v^* H_j v$  for some  $v$  such that  $v^* v = 1$ . Then

$$\begin{aligned} \langle \hat{x}, y \rangle &= \sum \hat{x}_j (v^* H_j v) \\ &\geq \lambda_{\min}(\sum \hat{x}_j H_j) \\ &= \langle \hat{x}, x \rangle \quad \text{for some } x \in \nabla_1 \\ &\geq \|\hat{x}\|^2 \quad \text{since } \hat{x} = \min(\text{co}\nabla_1) \end{aligned}$$

Therefore  $\|\hat{x}\| \leq \|\min(\text{co}\nabla_2)\|$  and since  $\nabla_1 \subset \nabla_2$ ,  $\hat{x} = \min(\text{co}\nabla_2)$ .

**Theorem 3**

$$\text{co}\nabla_1 = \text{co}\nabla_2$$

**Proof**

It suffices to prove that  $\text{dist}(x, \text{co}\nabla_1) = \text{dist}(x, \text{co}\nabla_2)$  for all  $x \in \mathbb{R}^q$ . Fix  $x = (x_1, \dots, x_q) \in \mathbb{R}^q$ . Define  $\hat{\nabla}_1$  and  $\hat{\nabla}_2$  by replacing  $H_j$  by  $H_j - x_j I$  in the definitions of  $\nabla_1$  and  $\nabla_2$ , respectively. Then, by lemma 4,

$$\begin{aligned} \text{dist}(x, \text{co}\nabla_1) &= \|\min(\text{co}\hat{\nabla}_1)\|_2 = \|\min(\text{co}\hat{\nabla}_2)\|_2 \\ &= \text{dist}(x, \text{co}\nabla_2) \end{aligned}$$

This holds for any  $x \in \mathbb{R}^q$ , and so the proof is complete.

In order to make effective use of  $\nabla_2$  in, say, an optimisation algorithm or a sensitivity calculation, there must be some reliable algorithm for computing  $\min(\text{co}\nabla_2)$  [or maybe  $\max(\text{co}\nabla_2)$  for sensitivity] given the  $\{M_j\}$ . The Appendix briefly describes an approach to this. The rest of this Section is concerned with properties of  $\nabla_2$  that will prove useful in the following Section.

Suppose that  $H_j = H_j^* \in \mathcal{M}(r)$ ,  $j = 1, 2, \dots, q$ , and let  $f: \mathbb{C}^r \rightarrow \mathbb{R}^q$  be defined by  $f_j(x) = x^* H_j x$  for  $x \in \mathbb{C}^r$ . Let

$$P^n = \{x \in \mathbb{C}^n \mid x^* x = 1\} \subset \mathbb{C}^n$$

and

$$S^n = \{x \in \mathbb{R}^{n+1} \mid x^T x = 1\} \subset \mathbb{R}^{n+1}$$

Note that this notation is not standard.

**Theorem 4**

For  $r = 2$ , there exists an affine map  $g: \mathbb{R}^3 \rightarrow \mathbb{R}^q$  such that  $f(P^2) = g(S^2)$ .

**Proof**  
Suppose

$$H_j = \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix}$$

for some  $j$ . Then

$$\begin{aligned} & [\cos\theta \quad \sin\theta(\cos\psi - i\sin\psi)] \begin{bmatrix} a & b \\ \bar{b} & c \end{bmatrix} x \\ & \begin{bmatrix} \cos\theta \\ \sin\theta(\cos\psi + i\sin\psi) \end{bmatrix} \\ & = a\cos^2\theta + c\sin^2\theta + 2\text{Re}(b\cos\theta \sin\theta(\cos\psi + i\sin\psi)) \\ & = (1/2)(a+c) + (1/2)(a-c)\cos 2\theta \\ & \quad + \sin 2\theta \text{Re}(b(\cos\psi + i\sin\psi)) \\ & = (1/2)(a+c) \end{aligned}$$

$$+ [(1/2)(a-c) \quad \text{Re}(b) \quad \text{Im}(b)] \begin{bmatrix} \cos 2\theta \\ \sin 2\theta \cos \psi \\ \sin 2\theta \sin \psi \end{bmatrix}$$

Let  $g_1: \mathbb{R}^3 \rightarrow \mathbb{R}$  be the affine map defined by the scalar and vector in the last equation so that  $f_1(P^2) = g_1(S^2)$  as shown. Define  $g_j$  for  $j = 2, 3, \dots, q$  similarly so that the resultant  $g: \mathbb{R}^3 \rightarrow \mathbb{R}^q$  is affine. Then  $f(P^2) = g(S^2)$  as desired.

**Corollary**

For  $f: \mathbb{C}^2 \rightarrow \mathbb{R}^q$  defined above, if  $q = 1$  or  $2$ , then  $f(P^2)$  is convex.

**Proof**

If  $q = 1$  or  $2$ , then for any affine map  $g: \mathbb{R}^3 \rightarrow \mathbb{R}^q$ ,  $g(S^2)$  is convex. Apply the theorem.

**Corollary**

Let  $f: \mathbb{C}^r \rightarrow \mathbb{R}^q$  be defined as above for arbitrary  $r \in \mathbb{Z}^+$ . If  $q = 1$  or  $2$ , then  $f(P^r)$  is convex.

**Proof**

Let  $x, y \in \mathbb{C}^r$ ,  $x \neq y$  and  $\|x\|_2 = \|y\|_2 = 1$ . It suffices to show that  $z(t) = tf(x) + (1-t)f(y) \in f(P^r)$  for all  $t \in [0, 1]$ . Let

$$\hat{H}_j = \begin{bmatrix} x^* \\ y^* \end{bmatrix} H_j \begin{bmatrix} x & y \end{bmatrix}$$

for all  $j = 1, 2, \dots, q$ , and let  $\hat{f}: \mathbb{C}^2 \rightarrow \mathbb{R}^q$  be defined as before in terms of the  $\hat{H}_j$ . Then  $f(x), f(y) \in \hat{f}(P^2) \subset f(P^r)$  and  $\hat{f}(P^2)$  is convex by the previous corollary. Thus  $z(t) \in f(P^2) \subset f(P^r)$  for all  $t \in [0, 1]$ , and so  $f(P^r)$  is convex.

This last corollary implies immediately that, if  $M: \Omega \rightarrow \mathcal{M}(K)$  for  $\Omega \in \mathbb{R}^q$  and  $q = 1$  or  $2$ , then the  $\nabla_2$  defined as before is convex. This will prove quite useful in the following Section.

Note that the results of this Section and Appendix could be used to study the sensitivity of clustered singular values to infinitesimal parameter variations. The aim of this paper, however, is to develop techniques for handling large perturbations, and so the following Section will make use of these results to compute the function  $\mu$ .

## 6 Computation of $\mu$

This Section will combine the results of the preceding two Sections and concentrate on the right-hand side inequality in property (k) of Section 4. Suppose in this Section that in the  $2n$ -tuple  $\mathcal{K}$ ,  $m_j = 1$  for each  $j$  (there are no repeated blocks). Then reduce  $\mathcal{K}$  to an  $n$ -tuple  $\mathcal{K} = (k_1, \dots, k_n)$  and let  $K = \Sigma k_j$ . Define  $D: \mathbb{R}^n \rightarrow \mathcal{M}(K)$  by

$$D(x) = \text{diag}(\exp(x_1)I, \exp(x_2)I, \dots, \exp(x_n)I)$$

For  $M_0 \in \mathcal{M}(K)$ , define  $M: \mathbb{R}^n \rightarrow \mathcal{M}(K)$  by

$$M(x) = D(x)M_0(D(x))^{-1} = D(x)M_0D(-x)$$

and let  $\sigma_{\max}(x)$ ,  $U_1$ ,  $V_1$ ,  $\nabla_1$ ,  $\nabla_2$  etc. be defined as in Section 4 for  $M(x)$ . Let  $A = U_1$  and  $B = U_2$  and assume that  $M$  is normalised so that  $\sigma_{\max}(0) = 1$ . Then write

$$A = [a_1 \ a_2 \ \dots \ a_r] = \begin{bmatrix} \alpha_1^* \\ \alpha_2^* \\ \vdots \\ \alpha_n^* \end{bmatrix}$$

and similarly for  $B$  in terms of  $\{b_j\}$  and  $\{\beta_j\}$ . Here  $\alpha_j^* \in \mathbb{C}^{K \times 1}$ ,  $\alpha_j^* \in \mathbb{C}^{K_j \times r}$  (similarly for  $B$ ). It is then easy to verify that  $H_j = \alpha_j \alpha_j^* - \beta_j \beta_j^*$ .

### Theorem 5

$\mu(M) = 1$  if, and only if,  $0 \in \nabla_2(M)$ .

### Proof

$0 \in \nabla_2(M)$  if, and only if:

(i) there exists  $y \in p^r$  such that

$$y^* H_j y = 0 \quad \text{for all } j$$

if, and only if,

(ii)  $\|\beta_j^* y\| = \|\alpha_j^* y\|$  for all  $j$

if, and only if,

(iii) for all  $j$ , there exists  $U_j \in U(k_j)$  such that

$$\beta_j^* y = U_j \alpha_j^* y$$

if, and only if,

(iv) there exists  $U \in \mathcal{U}$  such that  $By = UAy$   
( $U = \text{diag}(U_1, U_2, \dots, U_n)$ )

if, and only if,

(v)  $B^* U A y = y$

if, and only if,

(vi) there exists  $0 \neq x \in \mathbb{C}^K$  such that

$$AB^* U x = x$$

if, and only if,

(vii)  $M U x = x$  (since  $\sigma_{\max}(M) = 1$ )

if, and only if,

(viii)  $\rho(MU) = 1$

Therefore, using property (k) from Section 4,  $\mu(M) = 1$  if, and only if,  $0 \in \nabla_2(M)$ .

### Theorem 6

For  $n \leq 3$ ,  $\sigma_{\max}(M) = \mu(M)$  if, and only if,  $0 \in \text{co}\nabla_2$ .

### Proof

For any  $n$ ,  $\nabla_2$  depends only on  $n-1$  variables, since  $\nabla_2$  lies in the  $(n-1)$ -dimensional subspace of  $\mathbb{R}^n$  orthogonal to  $(1, 1, \dots, 1)$ ; i.e.  $D(x)M_0D(-x)$  is constant along any line parallel to the line through 0 and  $(1, 1, \dots, 1)$ . For  $n \leq 3$ , by the second corollary to theorem 3,  $\nabla_2 = \text{co}\nabla_2$ , and the result then follows from Theorem 1.

These two theorems have important implications for computing  $\mu$ . Suppose that, conceptually, a gradient search is to be used to compute

$$\inf_{D \in \mathcal{D}} \sigma_{\max}(DMD^{-1})$$

with  $\min(\text{co}\nabla_2)$  serving as a gradient when there are clustered singular values. This search can proceed until a local minimum is found, i.e.  $0 \in \text{co}\nabla_2$ . [As was mentioned earlier,  $\sigma_{\max}(DMD^{-1})$  is convex in  $D$ , and so such a local minimum would be global.] If, in fact,  $0 \in \nabla_2$  at this point, then, by theorem 5,  $\mu$  has also been found. Unfortunately, it is not generally true that  $0 \in \text{co}\nabla_2$  implies  $0 \in \nabla_2$ .

Theorem 6 implies that, for the case  $n \leq 3$  (three or fewer blocks),  $\nabla_2 = \text{co}\nabla_2$ , and so such a gradient search would (at least conceptually) always yield  $\mu$ . This result is rather remarkable since the optimisation involves only  $n-1$  variables, regardless of the size of the blocks. Furthermore, in view of theorem 3 and its corollaries, the computation of the generalised gradients for these cases should be quite straightforward. This provides the desired technique for exact analysis of multivariable systems with simultaneous input/output perturbations (i.e. two blocks), as well as more general situations.

## 7 Examples and computational experience

Two numerical examples and a discussion of computational experience are contained in this Section. The examples are intended to be purely illustrative, and no particular significance should be attributed to them.

The first example uses a  $3 \times 3$  matrix for which  $\mu$  will be computed for five different structured perturbations. The nominal matrix is

$$M = \begin{bmatrix} 4-2i & -\frac{1}{2}-\frac{1}{2}i & -10 \\ -24+6i & 3i & 60-80i \\ -\frac{6}{5} & -\frac{1}{5}-\frac{1}{5}i & 2+2i \end{bmatrix}$$

and the five uncertainty structures are shown in Table 1. The perturbation in case 1 is block diagonal, and so  $\mu$  can be computed directly from  $M$ . In cases 2–5, however, the matrix must be rearranged to make the perturbation block diagonal. This can always be done by pre- and postmultiplying  $M$  by appropriate matrices, denoted here by  $L$  and  $R$ , respectively. The last two columns in Table 1 give  $\mathcal{K}$  for that case and the corresponding  $\mu$  (LMR).

Cases 1–3 have three or fewer blocks, and  $\mu$  was computed using the approach suggested in Section 6. Cases 4 and 5 have more than three blocks, and  $\mu$  was computed using an algorithm based on theorems 1 and 2. The global maximum was found for case 4, as was verified by computing  $\inf \sigma_{\max}(DLMRD^{-1})$  and obtaining agreement to within the accuracy of the algorithms' termination conditions (in this case, four digits for single precision).

This example illustrates that  $\mu$  depends heavily on the assumed structure of the uncertainty. This should not be interpreted as a limitation on its usefulness. On the contrary, it emphasises the importance of having a method for analysing matrix perturbation problems that maintains structure. The next example illustrates how a particular uncertainty structure could naturally arise in a feedback system.

The following example was originally constructed by the author to illustrate that loop-at-a-time analysis was inadequate for studying simultaneous variations in multiloop systems and to introduce the use of singular values. An example was published by the author [6]. The actual design analysed here was

proposed by Sain *et al.* [15]. A thorough and lucid discussion of this example, including Sain's design, can be found in Reference 10.

Consider the feedback configuration in Fig. 1, where the nominal plant  $P$  has a transfer function

$$P(s) = \begin{bmatrix} \frac{9}{s+1} & \frac{-10}{s+1} \\ \frac{-8}{s+2} & \frac{9}{s+2} \end{bmatrix}$$

and

$$K(s) = \frac{1}{0.0159s} \begin{bmatrix} 9(s+1) & 10(s+2) \\ 8(s+1) & 9(s+2) \end{bmatrix}$$

The feedback compensation  $K$  is the product of the two elements of the compensator in References 15 and 10. Only the product  $K$  affects the system's feedback properties.

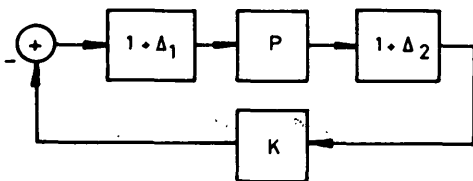


Fig. 1 Block diagram for example design

For this example, the nominal plant is considered to have simultaneous multiplicative perturbations [1] at the input and output. In applications, such perturbations might arise from actuators, sensors and unmodelled dynamics. The perturbations could then be weighted to reflect the fact that the level of uncertainty varies in frequency and direction. This example has no physical motivation, and so for simplicity this system will be evaluated for robustness with respect to unweighted size of the norm-bounded perturbations. It is desired to compute the smallest norm-bounded perturbation (as a function of frequency) which produces instability.

If only one perturbation is considered at a time, then  $\mu$  is simply the maximum singular value of the transfer function  $(I + PK)^{-1}PK$  or  $(I + KP)^{-1}KP$ , for an output or input perturbation, respectively. These two transfer functions happen to be the same for this example, and their maximum singular value is plotted as the lower curve in Fig. 2. This implies, for example, that there exists a destabilising pertur-

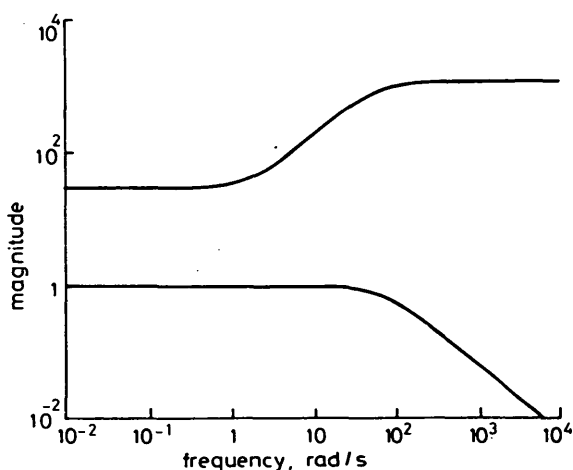


Fig. 2 Singular-value bounds

bation of norm 1 and that all smaller perturbations can be tolerated without instability. Unfortunately, this provides very little information (other than the upper bound) on the tolerable level for simultaneous perturbations; i.e. these maximum singular values are a lower bound for  $\mu$  for simultaneous perturbations.

In order to analyse the system for simultaneous variations, it is rearranged to isolate the  $\Delta$ 's as a block-diagonal perturbation in standard feedback configuration, as in Fig. 3. For this example, one choice is

$$M = \begin{bmatrix} (I + KP)^{-1}KP & (I + KP)^{-1}K \\ -(I + PK)^{-1}P & (I + PK)^{-1}PK \end{bmatrix}$$

with

$$\Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}$$

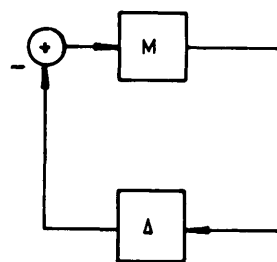


Fig. 3 Standard feedback configuration

If the block structure were ignored at this point, then  $\sigma_{\max}(M)$  (the upper curve in Fig. 2) would provide a tight bound on tolerable perturbations  $\Delta$ . Unfortunately,  $\sigma_{\max}(M)$  provides only a conservative bound for the block-diagonal case.

For the structure as given, the corresponding  $\mu$  is plotted in Fig. 4. This plot has several interpretations. The simplest is that the system can be stabilised by simultaneous perturbations with norms approximately equal to 0.1, and smaller perturbations may be tolerated without instability. The bounds in Fig. 2 were almost useless for this example. More generally, it is quite easy to construct examples where the gap for both bounds is arbitrarily large.

The relatively large value of  $\mu$  in this analysis should not be considered as an indictment of this design. It merely indicates that the design has poor margins relative to this particular uncertainty structure; with respect to another structure they may be much better (for example, the margins for uncertainty on just one side may be good). A design that tolerated larger simultaneous variations of the type considered here

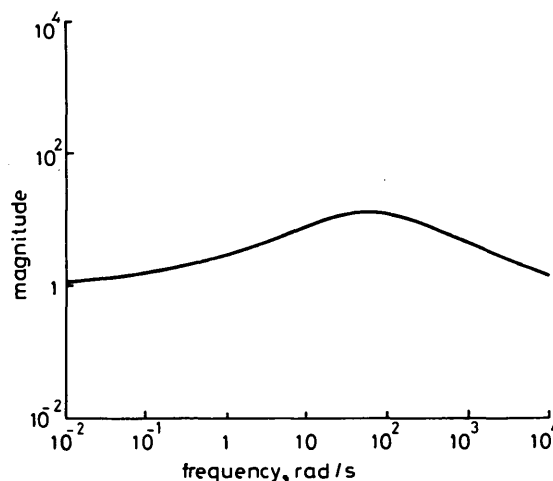


Fig. 4 Plot of  $\mu$  for example design

Table 1: Uncertainty structures

Case	$\Delta$	$L$	$R$	$\mathcal{K}$	$\mu$
1	$\begin{bmatrix} \Delta_1 & 0 & 0 \\ 0 & \Delta_2 & 0 \\ 0 & 0 & \Delta_3 \end{bmatrix}$	$I_3$	$I_3$	$(1, 1, 1, 1, 1, 1)$	10.24
2	$\begin{bmatrix} \Delta_1 & 0 & 0 \\ 0 & 0 & \Delta_2 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$	$(1, 1, 1, 1)$	4.65
3	$\begin{bmatrix} \Delta_1 & 0 & 0 \\ \Delta_2 & & \Delta_3 \\ 0 & & \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$(1, 1, 1, 1, 1, 2)$	102.8
4	$\begin{bmatrix} \Delta_1 & 0 & 0 \\ \Delta_2 & \Delta_3 & 0 \\ 0 & 0 & \Delta_4 \end{bmatrix}$	same	same	$(1, 1, 1, 1, 1, 1, 1)$	10.8
5	$\begin{bmatrix} \Delta_1 & 0 & 0 \\ \Delta_2 & \Delta_3 & 0 \\ 0 & 0 & \Delta_3 \end{bmatrix}$	same	same	$(1, 1, 2, 1, 1, 1)$	10.6

would have to sacrifice some other aspect of performance. There was no physical basis for the uncertainty: it was chosen to be illustrative. Thus the results should not be interpreted too broadly.

This example reiterates the point that vastly different numbers are obtained, depending on the assumed structure of the problem. For practical problems, the structure is dictated to a large degree by physical reality and engineering constraints. The extent to which a design engineer can capture and handle the natural structure determines the extent to which the conclusions based on any analysis are relevant to the practical problem. It is hoped that the ideas introduced in this paper and illustrated in the examples will make a contribution towards better techniques for handling structured uncertainty in feedback systems.

An important consideration in the application of  $\mu$  is a numerical software. In as much as the ideas in this paper are relatively new and have not been published before, it may be some time before reliable numerical software (as in Linpack or Eispack) is available to compute  $\mu$ . The computational experience to date is most encouraging, however. Programs have been developed to compute both bounds in property (k) of Section 4. Recall that the lower bound is always an equality (see Section 4), but the global maximum may be difficult to find. The optimisation problem in the upper bound is convex, but it is only guaranteed to yield  $\mu$  for three or fewer blocks (see Section 6).

The ratio between the lower and upper bounds produced by these new programs has been computed for over 50 000 pseudo-randomly generated matrices of dimension three to ten, mostly with scalar blocks. The worst-case ratio was approximately 0.95, although examples have been constructed analytically where the ratio was  $\approx 0.85$ . It is interesting to note that the ratio seems not to decrease after four dimensions.

This is suggestive but, of course, not conclusive. As expected, for three dimensions it is always 1 to within reasonable numerical error.

Because the computer programs are experimental, i.e. containing many diagnostics and obviously inefficient code, it is impossible to draw any meaningful conclusions about computational speed. As a single data point, it took approximately three times longer to compute the curve in Fig. 4 as it did to compute the bounds in Fig. 2.

## 8 Summary and conclusions

This paper has introduced a general approach for analysing linear systems with structured uncertainties based on a new generalised spectral theory for matrices. This basic theory addresses the norm-bounded perturbation problem with arbitrary structure. The strongest results are for perturbations with three or fewer blocks, for which (conceptual) algorithms with guaranteed convergence were proposed. One application of these results is a generalisation of standard singular-value analysis techniques for multivariable feedback systems to treat simultaneous input/output uncertainty (i.e. two blocks).

These results are merely a beginning, and much more work remains to be done. For example, existing multivariable control methods provide little more than minor extensions of SISO techniques. Initial study indicates that consideration of multiple simultaneous perturbations leads to wholly new phenomena, the explanation of which will provide a far deeper understanding of multiloop feedback systems. Linear multivariable control theory will need to be thoroughly re-examined in this new light. It is hoped that the more general results in this paper could also provide the beginning of a nontrivial theory of decentralised control and/or large-scale systems, where use of structural information is essential.



## 9 Acknowledgments

Many people contributed to this paper, but I would particularly like to thank Dr. Joe Wall, of Honeywell Systems and Research Center, for his help throughout the research and writing. Jim Freudenberg of the University of Illinois and Honeywell SRC made a major contribution to the implementation of the algorithms used in the examples and offered many useful suggestions regarding the paper.

This work has been supported by Honeywell internal research and development funding, the US Office of Naval Research under ONR research grant N00014-82-C-0157, and the US Air Force Office of Scientific Research grant F49620-82-C-0090.

This work is in the public domain in the USA.

## 10 References

- 1 DOYLE, J.C., and STEIN, G.: 'Multivariable feedback design: Concepts for a classical/modern synthesis', *IEEE Trans.*, 1981, AC-26, pp. 4-16
- 2 POSTELTHWAITE, I., EDMUNDS, J.M., and MacFARLANE, A.G.J.: 'Principal gains and principal phases in the analysis of linear multivariable feedback systems', *ibid.*, 1981, AC-26, pp. 32-46
- 3 SAFONOV, M.G., LAUB, A.J., and HARTMANN, 'Feedback properties of multivariable systems: The role and use of the return difference matrix', *ibid.*, 1981, AC-26, pp. 47-65
- 4 CRUZ, J.B., FREUDENBERG, J.S., and LOOZE, D.P.: 'A relationship between sensitivity and stability of multivariable feedback systems', *ibid.*, 1981, AC-26, pp. 66-74
- 5 LEHTOMAKI, N.A., SANDELL, N.R., Jr., and ATHANS, M.: 'Robustness results in linear-quadratic Gaussian based multivariable control designs', *ibid.*, 1981, AC-26, pp. 75-92
- 6 DOYLE, J.C.: 'Robustness of multiloop linear feedback systems' 17th IEEE conference on decision and control, San Diego, USA, Jan. 1979
- 7 DOYLE, J.C.: 'Multivariable design techniques based on singular value generalisation of classical control' AGARD lecture series 117 on multivariable analysis and design techniques. Sept. 1981
- 8 WALL, J.E., DOYLE, J.C., and HARVEY, C.A.: 'Tradeoffs in the design of multivariable feedback systems'. Proceedings of 18th Allerton conference on communication control and computing, Oct. 1980, pp. 715-725
- 9 DOYLE, J.C.: 'Limitations on achievable performance of multivariable feedback systems'. AGARD lecture series 117 on multivariable analysis and design techniques, September. 1981
- 10 FREUDENBERG, J.S., LOOZE, D.P., and CRUZ, J.B.: 'Robustness analysis using singular value sensitivities', *Int. J. Control*, 1982, pp. 95-116
- 11 KATO, T.: 'Perturbation theory for linear operators' (Springer-Verlag, 1976)

- 12 WILKINSON, J.H.: 'The algebraic eigenvalue problem' (Clarendon Press, 1965)
- 13 RELICH, F.: 'Perturbation theory of eigenvalue problems' (Gordon & Breach, 1969)
- 14 LUENBERGER, D.G.: 'Introduction to linear and nonlinear programming' (Addison-Wesley, 1973)
- 15 SAIN, M.K., MA, A., and PERKINS, D.: 'Sensitivity issues in decoupled control system design'. Proceedings of Southeast symposium on system theory

## 11 Appendix

Let  $\{H_j\}_{j=1}^q$ ,  $f$ , and  $P^r$  be defined as in the paragraph preceding theorem 3, and let  $\nabla_2 = f(P^r)$ . To compute  $\min(\text{co}\nabla_2)$ , consider the following algorithm to produce a sequence  $\{x_n\}$ , where  $x_n = [x_n^1, x_n^2, \dots, x_n^q] \in \mathbb{R}^q$ :

- (a) Pick any  $v_1 \in P^r$  and let  $x_1 = f(v_1)$ .
- (b) Inductively define  $x_n = \min(\text{co}[x_{n-1}, f(v_{n-1})])$ .
- (c) Then  $v_n$  is any unit eigenvector for  $\lambda_{\min}(\Sigma x_n^T H_j)$ .

This produces a sequence  $\{x_n\}$  that has nonincreasing norm, and thus  $\{\|x_n\|\}$  converges. Therefore there exists a subsequence  $\{x_{n_k}\}$  such that both  $\{x_{n_k}\}$  and  $\{v_{n_k}\}$  converge to, say,  $x_0$  and  $v_0$ , respectively. Note that  $y = f(v_n)$  minimises the  $\min\langle x_n, y \rangle$  and  $\langle x_n, f(v_n) \rangle = \lambda_{\min}(\Sigma x_n^T H_j)$ . Then, by taking limits,  $\langle x_0, f(v_0) \rangle = \lambda_{\min}(\Sigma x_0^T H_j)$  and  $y = f(v_0)$  minimises  $\min\langle x_0, y \rangle$ .

Now suppose  $x_0 \neq \min(\text{co}\nabla_2)$  (leading to a contradiction). Then by definition of the convex hull,  $\langle x_0, f(v_0) \rangle < \|x_0\|^2$  and  $\|x_0\| - \|\min(\text{co}[x_0, f(v_0)])\| > \epsilon$  for some  $\epsilon > 0$ . By existence of a convergent subsequence, there exists  $N$  sufficiently large such that  $\|x_N - x_0\| < \epsilon/2$  and  $\|f(v_N) - f(v_0)\| < \epsilon/2$ . Then

$$\begin{aligned} \|x_0\| &> \epsilon + \|\min(\text{co}[x_0, f(v_0)])\| \\ &\geq \epsilon + [\min(\text{co}[x_N, f(v_N)]) - \epsilon/2] \\ &= \epsilon/2 + \|x_{N+1}\| \end{aligned}$$

This contradicts the fact that  $\{\|x_n\|\}$  is nonincreasing, and thus  $\{x_n\}$  converges to  $\min(\text{co}\nabla_2)$ .

In this algorithm  $f(v_n)$  is an intersection point of  $\nabla_2$  and the support hyperplane perpendicular to  $x_n$ . There are, of course, a number of other ways that this point could be used to obtain a convergent algorithm. The particular algorithm outlined here is rather crude in not taking full advantage of the structure of  $\nabla_2$ .